

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 20:49:12 ; Search time 1540.85 Seconds
(without 21 elements)

8229.920 Million cell updates/sec

Title: US-09-988-971-1_COPY_415_1197
Page: 783

Sequence: 1 atgggaagtcctgccagcag.....ctgtctctctgcatgagcc 783

Scoring table: IDENTITY_NUC

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
```

Listing first 45 summaries

Database :

```

1: em_eatba:*
2: em_eatbm:*
3: em_eatm:*
4: em_eatnu:*
5: em_eatv:*
6: em_eatvp:*
7: em_eatzo:*
8: em_hc:*
9: gb_eatl:*
10: gb_eat2:*
11: gb_hc:*
12: gb_gsc:*
13: gb_gsc2:*
14: gb_gsc5:*
15: em_gsfun:*
16: em_escom:*
17: gb_gsr:*
18: em_gsr_hm:*
19: em_gsr_inv:*
20: em_gsr_pin:*
21: em_gsr_vt:*
22: em_gsr_fun:*
23: em_gsr_nms:*
24: em_gsr_nms2:*
25: em_gsr_ohes:*
26: em_gsr_pro:*
27: em_gsr_rod:*

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	734.6	93.8	1002	14	B0052308	B0052308 AGENCOURT
2	664	83.5	B0054281		AGENCOURT	B0054281 AGENCOURT
3	662	82.0	1069	14	B0052468	B0052468 AGENCOURT
4	488.6	62.4	877	9	AL541041	AL541041 AL541041
5	477	60.9	526	11	AKO20837	AKO20837 Mus muscu
6	402	51.3	566	12	BG284179	BG284179 602408222

7	386.2	49.3	986	14	B0054285	AGENCOUNT
8	331.2	48.3	878	16	B0053486	AGENCOUNT
9	326.8	41.7	778	12	BG178487	BG178487 602328305
10	319.4	40.8	577	9	AL844311	AL844311 AL844311
11	314.4	40.2	614	9	AL844307	AL844307 AL844307
12	311.6	39.8	606	9	AL844309	AL844309 AL844309
13	286.2	37.8	660	10	B8635615	B8635615 B8635615
14	283.6	36.2	377	9	AA959151	AA959151 v251906.r
15	275.6	35.2	627	10	B8619854	B8619854 B8619854
16	258.4	33.0	794	12	BG677557	BG677557 602624102
17	220.6	28.2	603	14	B0553005	B0553005 HA019E02-
18	220	28.1	670	9	AT510095	AT510095 m13c04.y
19	215.8	27.6	640	14	B0553006	B0553006 HA019E02-
20	163.8	20.9	642	9	AL844308	AL844308 AL844308
21	162.2	20.7	871	14	BQ436113	BQ436113 AGENCOUNT
22	160.6	20.5	570	9	AT471720	AT471720 ta16a01.x
23	160.2	20.5	909	9	AL549826	AL549826 AL549826
24	160.2	20.5	958	9	AL539427	AL539427 AL539427
25	159.8	20.4	902	9	AL551370	AL551370 AL551370
26	159.2	20.3	1063	14	BQ072745	BQ072745 AGENCOUNT
27	155	20.3	972	14	BQ707614	BQ707614 AGENCOUNT
28	158.8	20.3	619	9	AL844312	AL844312 AL844312
29	151.8	19.4	849	13	B1769183	B1769183 603053793
30	149.8	19.1	541	13	B1889899	B1889899 480839 MA
31	136.6	17.4	389	10	BEO15229	BEO15229 127457 MA
32	135.4	17.3	611	9	AL844310	AL844310 AL844310
33	135.6	17.2	673	10	B8638252	B8638252 B8638252
34	133.6	17.1	723	14	BM950089	BM950089 UT-M-EHDP
35	131	16.7	526	10	BB283748	BB283748 BB283748
36	124.8	15.9	2814	11	BC003253	BC003253 Mus muscu
37	123.6	15.8	773	9	AA453694	AA453694 AA453694
38	119.8	15.3	999	10	BB610404	BB610404 BB610404
39	119.4	15.2	976	9	AL543079	AL543079 AL543079
40	119	15.2	591	13	B1345674	B1345674 374563 MA
41	119	15.2	789	9	AUI34909	AUI34909 AUI34909
42	118.4	15.1	671	13	B1333488	B1333488 ppr1n.pk0
43	118	15.1	541	10	BB709137	BB709137 BB709137
44	117.2	15.0	550	9	AT330952	AT330952 m13c04.x
45	115.4	14.7	546	10	BB397325	BB397325 BB397325

ALIGNMENTS

RESULT 1
BQ052308

LOCUS	BQ052308	1002 bp	mRNA	linear	EST 29-MAR-2002
DEFINITION	AGENCOURT_6868571 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:59335424				

ACCESSION BQ052308

KEYWORDS EST.

ORGANISM Homo

1

REFERENCE 1 (bases 1 to 1002)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNALS

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clonal distribution: MCC clonal distribution information

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>

Plate: LLCM2118 row: d column: 23
High quality sequence from: 670

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT      221 a      296 c      288 g      197 t
ORIGIN

```

```

Query Match      93.8%; Score 734.6; DB 14; Length 1002;
Best Local Similarity 98.3%; Pred. No. 1.5e-177;
Matches 774; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

```

```

QY 1 ATGGAGAGCTGCGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGC 60
DB 96 ATGGAGAGCTGCGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGC 155
QY 61 CAAGGCCAGGACCTGTGACATGAAAGCAGAGAACCAAGCCACAGCCCTGCGCTG 120
DB 156 CAAGGCCAGGACCTGTGACATGAAAGCAGAGAACCAAGCCACAGCCCTGCGCTG 215
QY 121 GGCAGTTTCCCGCAGTGGCCCGCCGACGCTGTGAGAGCTCGGGAGCCATTGACC 180
DB 216 GGCAGTTTCCCGCAGTGGCCCGCCGACGCTGTGAGAGCTCGGGAGCCATTGACC 275
QY 181 ATGCTCTCTGAGAGATGAGATGCTGTCGCTGTCTGAAGTCTCAAGCAGAGATAT 240
DB 276 ATGCTCTCTGAGAGATGAGATGCTGTCGCTGTCTGAAGTCTCAAGCAGAGATAT 335
QY 241 AACATCCCAAGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGGCTGTATGAGGCGCTGAGC 300
DB 336 AACATCCCAAGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGGCTGTATGAGGCGCTGAGC 395
QY 301 AGGAGAAAGCAGAGAACTGCTGTTTAACTCTGGAAACCTTGGAGGGGCTTCTCTATC 360
DB 396 AGGAGAAAGCAGAGAACTGCTGTTTAACTCTGGAAACCTTGGAGGGGCTTCTCTATC 455
QY 361 CGGAGAGACCAAGCAGAGAGGCTTTACTCTGTGACGTCCGCTCAGCCGCTTGA 420
DB 456 CGGAGAGACCAAGCAGAGAGGCTTTACTCTGTGACGTCCGCTCAGCCGCTTGA 515
QY 421 TCTGGAGACCGGATCAGACCTACAGATCACTGCTTGAACAATGGCTGCTGATCATC 480
DB 516 TCTGGAGACCGGATCAGACCTACAGATCACTGCTTGAACAATGGCTGCTGATCATC 575
QY 481 TCACCGGCTCCTCACTTCCCTCACTCAGAGCCCTGTGAGCACTTACTCTGAGTGGCG 540
DB 576 TCACCGGCTCCTCACTTCCCTCACTCAGAGCCCTGTGAGCACTTACTCTGAGTGGCG 635
QY 541 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTGCAAGAGGCTGCGCCCTCTCT 600
DB 636 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTGCAAGAGGCTGCGCCCTCTCT 695
QY 601 GGCAGAGATATACCCCTTACTGTGACTGTGACAGAGAACCACTCAACTGGAAGAGCTG 660
DB 696 GGCAGAGATATACCCCTTACTGTGACTGTGACAGAGAACCACTCAACTGGAAGAGCTG 755
QY 661 GACAGCTCCTCTCTGTTTCTGAAGCTGCCACA-GGGAGAGATCTTCTCTCAAGTGAAGG 719
DB 756 GACAGCTCCTCTCTGTTTCTGAAGCTGCCACAAGAGAGAGATCTTCTCTCAAGTGAAGG 815
QY 720 --TCTCCGAGAGTCCCTCAGCTTCAATCAGCTGAGATGACGA-GGCTGTCTCTTGA 776
DB 816 GTCTCGGGGAGTCCCTCAGCTTCAATCAGCTGAGATGACGAAGGCTGTCTCTTGA 875

```

```

QY 777 TGATGCC 783
DB 876 TGATGCC 882

```

```

RESULT 2
B0054281 1020 bp mRNA linear EST 29-MAR-2002
AGENCOURT 6810234 NIH-MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
B0054281
B0054281.1 GI:19813621
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1020)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LNCM2125 row: j column: 11
High quality sequence stop: 556.
Location/Qualifiers
1. 1020
source

```

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/clone_lib="NIH-MGC_106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT      219 a      311 c      283 g      204 t      3 others
ORIGIN

```

```

Query Match      83.5%; Score 654; DB 14; Length 1020;
Best Local Similarity 95.4%; Pred. No. 6.6e-157;
Matches 748; Conservative 0; Mismatches 27; Indels 9; Gaps 7;

```

```

QY 1 ATGGAGAGTCCGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGC 60
DB 212 ATGGAGAGTCCGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGC 271
QY 61 CAAGGCCAGGACCTGTGACATGAAAGCAGAGAACCAAGCCACAGCCCTGCGCTG 120
DB 272 CAAGGCCAGGACCTGTGACATGAAAGCAGAGAACCAAGCCACAGCCCTGCGCTG 331
QY 121 GGCAGTTTCCCGCAGTGGCCCGCCGACGCTGTGAGAGCTCGGGAGCCATTGACC 180
DB 332 GGCAGTTTCCCGCAGTGGCCCGCCGACGCTGTGAGAGCTCGGGAGCCATTGACC 391
QY 181 ATGCTCTCTGAGAGATGAGATGCTGTCGCTGTCTGAAGTCTCAGCAGAGATAT 240
DB 392 ATGCTCTCTGAGAGATGAGATGCTGTCGCTGTCTGAAGTCTCAGCAGAGATAT 451
QY 241 AACATCCCAAGCTCCAGCGTGGCCAAAGTCTCCCAATGGGTGGCTGTATGAGGCGCTGAGC 300

```

Db 452 AACATCCCAGCCTCCACGTCGCAAGATCTCCCATGGGTGGCTGTATGAGGGCTGAGC 511
 Qy 301 AGGGAAGAAAGCAAGAACTGCTGTTTACCTTGGGAACCTTGGAGGGGCTTCTCATC 360
 Db 512 AGGGAAGAAAGCAAGAACTGCTGTTTACCTTGGGAACCTTGGAGGGGCTTCTCATC 571
 Qy 361 CGGAGAGCCAGACAGAGAGGCTTCACTCTCTGTCAAGTCCGCTCAGCCGCTGCA 420
 Db 572 CGGAGAGCCAGACAGAGAGGCTTCACTCTCTGTCAAGTCCGCTCAGCCGCTGCA 631
 Qy 421 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGAACATGCTGGCTGTACATC 480
 Db 632 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGAACATGCTGGCTGTACATC 691
 Qy 481 TCACCGCGCTCACTTCCCTCACTCCAGG-CCCTGTGAGCACTTACTGTAGCTGAGC 539
 Db 692 TCACCGCGCTCACTTCCCTCACTCCAGG-CCCTGTGAGCACTTACTGTAGCTGAGC 751
 Qy 540 GGATGACATCTGCTGCTTCA-CTCAAGAGCCCTGTGTCTCTGCAAGGGCTGGCCCTCC 598
 Db 752 GGATGACATCTGCTGCTTCA-CTCAAGAGCCCTGTGTCTCTGCAAGGGCTGGCCCTCC 811
 Qy 599 CTGGCAAGATATA-CCCCCTACCTGTGACTGTGAGAGACACACTCACTGAGAAAG 657
 Db 812 CTGGCAAGATATA-CCCCCTACCTGTGACTGTGAGAGACACACTCACTGAGAAAG 871
 Qy 658 CTGGACAGCTCCCTCTG-TTTTCTGAAGCTGACAC-GGGAGAGAGCTCTTCTCAGT 714
 Db 872 CTGGACAGCTCCCTCTG-TTTTCTGAAGCTGACAC-GGGAGAGAGCTCTTCTCAGT 931
 Qy 715 GAGGG--TCTCCGGAGTCCCTCAGCTTCTACATCAGCTG--AATGACAGAGCTGTCT 771
 Db 932 GAGGGTCTTCCGGAGTCCCTCAGCTTCTACATCAGCTG--AATGACAGAGCTGTCT 991
 Qy 772 TTGG 775
 Db 992 TTGG 995

RESULT 3
 LOCUS B0052468 1069 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
 5' mRNA sequence.

ACCESSION B0052468
 VERSION B0052468
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (Bases 1 to 1069)

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L12M2118 row: n column: 13
 High quality sequence stop: 681.

FEATURES

Location/Qualifiers
 1..1069
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5933772"
 /clone_11b="NIH_MGC_106"
 /tissue_type="natural killer cells, cell line"

/lab_host="DH10B (phage-resistant)"
 /note="Organ: Blood; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC library."

BASE COUNT 230 a 328 c 300 g 205 t 6 others
 ORIGIN
 Query Match 82.0%; Score 642; DB 14; Length 1069;
 Best Local Similarity 93.4%; Pred. No. 8.1e-154;
 Matches 734; Conservative 0; Mismatches 0; Indels 52; Gaps 4;

Qy 1 ATGGGAAGTGTGCGCCAGCAAGAAATCTGTCCAAAGCCCAAGCTTGAATCTCTGTC 60
 Db 89 ATGGGAAGTGTGCGCCAGCAAGAAATCTGTCCAAAGCCCAAGCTTGAATCTCTGTC 148
 Qy 61 CAAGCCAGAGGACCTGTGACATGAGAAAGCAAGAGCAAGGACAGCCGTGGCCCTG 120
 Db 149 CAAGCCAGAGGACCTGTGACATGAGAAAGCAAGAGCAAGGACAGCCGTGGCCCTG 208
 Qy 121 GGCAGTTTCCCGAGAGTGGCCCGGAGACCTGTGAGACATCGGGAGGACATGAGC 180
 Db 209 GGCAGTTTCCCGAGAGTGGCCCGGAGACCTGTGAGACATCGGGAGGACATGAGC 268
 Qy 181 ATGCTCTGTAGATGAGACATGTGTGACCGGTGTGTCTGAAGTCTCAGCAGAGAGAT 240
 Db 269 ATGCTCTGTAGATGAGACATGTGTGACCGGTGTGTCTGAAGTCTCAGCAGAGAGAT 328
 Qy 241 AACATCCCAAGGCTCCAGGTGGCCAAAGTCTCCCAAGGTGGTGTGAGGGGCTGAGC 300
 Db 329 AACATCCCAAGGCTCCAGGTGGCCAAAGTCTCCCAAGGTGGTGTGAGGGGCTGAGC 388
 Qy 301 AGGGAAGAAAGCAAGAACTGCTGTGTTAATCTGTGAAACCTGTGAGGGGCTTCTCATC 360
 Db 389 AGGGAAGAAAGCAAGAACTGCTGTGTTAATCTGTGAAACCTGTGAGGGGCTTCTCATC 448
 Qy 361 CGGAGAGCCAGACAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
 Db 449 CGGAGAGCCAGACAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 508
 Qy 421 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGAACATGCTGGCTGTACATC 480
 Db 509 TCTTGGAGCCGGATCAGACATCAAGATCCATGCTTGAACATGCTGGCTGTACATC 568
 Qy 481 TCACCGCGCTCACTTCCCTCACTCCAGGCGCT-GGTGAACATTAATCTGAGCTGGC 539
 Db 569 TCACCGCGCTCACTTCCCTCACTCCAGGCGCTGGGTGAGCAATTAATCTGAG----- 623
 Qy 540 GGATGACATCTGCTGCTTCACTCAAGAGCCCTGTGTCTCTGAGAGGGCTGGCCCTCC 599
 Db 624 -----GGCTGTGGCCCTCC 639
 Qy 600 TGGCAAGATATACCTTACCTGATGAGTGGAGAGACACACTCACTGAGAAAGAGCT 659
 Db 640 TGGCAAGATATACCTTACCTGATGAGTGGAGAGACACACTCACTGAGAAAGAGCT 699
 Qy 660 GAGACGCTCCCTCTGTTTCTGAAGTGGACAC-GGGAGAGAGCTCTTCTGAGTGA 718
 Db 700 GAGACGCTCCCTCTGTTTCTGAGAGTGGACACAGAGGGAGAGAGTCTTCTGAGTGA 759
 Qy 719 GTCTCCGGAG-AGTCCCTGAGCTTCACTGAGCTGAAATGACAGAGGTGCTTTGGAT 777
 Db 760 GTCTCCGGAGAGTCCCTCACTGCTTCACTGAGCTGAAATGACAGAGGTGCTTTGGAT 819
 Qy 778 GATGCC 783
 Db 820 GATGCC 825

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>

Plate: L1AM10418 row: c column: 07
 High quality sequence start: 2
 High quality sequence stop: 566.
 Location/Qualifiers

FEATURES

source

1..566
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520382"
 /clone_lib="NIH MGC 91"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Prostate; Vector: PCMV-SPORE6; Site_1: NCI;
 Site_2: Sall; Cloned unidirectionally; oligo-dr primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."

BASE COUNT

116 a 187 c 152 g 110 t 1 others

ORIGIN

Query Match 51.3%; Score 402; DB 12; Length 566;
 Best Local Similarity 100.0%; Pred. No. 1.9e-92;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 GGCTTAACTCTGTGATCCGCTCAGCCGCTGATCTCGGAGCCGATCAGAC 441
 DB 13 GGCTTAACTCTGTGATCCGCTCAGCCGCTGATCTCGGAGCCGATCAGAC 72
 QY 442 TACGATCACTGCTGATGATGCTGCTGATCATCTGATCCGCTCAGCTTCCC 501
 DB 73 TACGATCACTGCTGATGATGCTGCTGATCATCTGATCCGCTCAGCTTCCC 132
 QY 502 TCACTCCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 561
 DB 133 TCACTCCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 192
 QY 562 AAGAGCCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 621
 DB 193 AAGAGCCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 252
 QY 622 GTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
 DB 253 GTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
 QY 682 GAAGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
 DB 313 GAAGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 372
 QY 742 TACATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
 DB 373 TACATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 414

RESULT 7

LOCUS B0054265 986 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
 5', mRNA sequence.

ACCESSION B0054265
 VERSION B0054265.1 GI:19813605
 KEYWORDS EST.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 986)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
 Plate: L1AM2125 row: i column: 12
 High quality sequence stop: 515.
 Location/Qualifiers

FEATURES

source

1..986
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5936339"
 /clone_lib="NIH MGC 106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dr priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGAGGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT

211 a 291 c 276 g 207 t 1 others

ORIGIN

Query Match 49.3%; Score 386.2; DB 14; Length 986;
 Best Local Similarity 98.0%; Pred. No. 2.7e-89;
 Matches 444; Conservative 0; Mismatches 3; Indels 6; Gaps 5;

QY 1 ATGGAGATGCTGCGCAGAGAAATCTTGTGCAAGCCCAAGTTGATCTCTGTC 60
 DB 279 ATGGAGATGCTGCGCAGAGAAATCTTGTGCAAGCCCAAGTTGATCTCTGTC 338
 QY 61 CAAGGCAAGGAGCTGTGACCATGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 120
 DB 339 CAAGGCAAGGAGCTGTGACCATGGAAGCAAGCAAGCAAGCAAGCAAGCAAG 398
 QY 121 GCGATTTCCGCGAGGTGCGCCGCGAGCTGTGCTGATGCTGGGAGCAATTGAC 180
 DB 399 GCGATTTCCGCGAGGTGCGCCGCGAGCTGTGCTGATGCTGGGAGCAATTGAC 458
 QY 399 GCGATTTCCGCGAGGTGCGCCGCGAGCTGTGCTGATGCTGGGAGCAATTGAC 458
 QY 181 ATGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 DB 459 ATGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 518
 QY 241 AACATCCCAAGCTGCAAGTGGCCAAAGTCTCCATGAGGTGCTGATGAGGCTGAGC 300
 DB 519 AACATCCCAAGCTGCAAGTGGCCAAAGTCTCCATGAGGTGCTGATGAGGCTGAGC 578
 QY 301 AGGAGAAAGAGAGAACTGCTGTTGTACTGAGGAACCTGAGAGGGG-CCTTCTCAT 359
 DB 579 AGGAGAAAGAGAGAACTGCTGTTGTACTGAGGAACCTGAGAGGGG-CCTTCTCAT 638
 QY 360 CCGGAGG-AGCCAAGCAGAGAGGCTTTTACTCTGCTCAT-CCGCTCAGCCG-CCC 416
 DB 639 CCGGAGGAGGAGAGAGGCTTTTACTCTGCTCAT-CCGCTCAGCCG-CCC 698
 QY 417 TGCATCTGGG--ACCGATCAGACATACAGG 447
 DB 699 TGCATCTGGGAGCCGAGATCAGACCTCAGAG 721

RESULT 8

LOCUS B0053486 878 bp mRNA linear EST 29-MAR-2002
 DEFINITION AGENCOURT_6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253
 5', mRNA sequence.
 ACCESSION B0053486

KEYWORDS	B0053486.1 GI:19812826
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE	1 (bases 1 to 878)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabbs-remail.nih.gov
	Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
	cDNA Library Preparation: Rubin Laboratory
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
	DNA Sequencing by: Agencourt Bioscience Corporation
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
	Plate: LNCM2122 row: 1 column: 06
	High quality sequence stop: 394.
FEATURES	Location/Qualifiers
source	1..878
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGC:593253"
	/clone_lib="NIH-MGC_106"
	/tissue_type="natural killer cells, cell line"
	/lab_host="DH10B (phage-resistant)"
	/note="Organ: blood; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."
BASE COUNT	201 a 253 c 233 g 190 t 1 others
ORIGIN	
Query Match	42.3%; Score 331.2; DB 14; Length 878;
Best Local Similarity	93.1%; Pred. No. 3.2e-74;
Matches 390; Conservative	0; Mismatches 24; Indels 5; Gaps 4.
Qy	1 ATGGGAAGTCTGCCACGACGAAAGAAATCTCTGCCAAGCCCAAGTTAGATTCTCTGTC 60
Db	273 ATGGGAAGTCTGCCACGACGAAAGAAATCTCTGCCAAGCCCAAGTTAGATTCTCTGTC 332
Qy	61 CAAGGCCAGGACCTGTGACCATGAAAGAGAGAGAAAGCAAGGCCACGCGAGGCGCTG 120
Db	333 CAAGGCCAGGACCTGTGACCATGAAAGAGAGAGAAAGCAAGGCCACGCGAGGCGCTG 392
Qy	121 GGCAGATTTCCCGCAGGTGGCCGCCGCGAGCTGTGCTGAGACTCGGAGGACCATTTGAC 180
Db	393 GGCAGATTTCCCGCAGGTGGCCGCCGCGAGCTGTGCTGAGACTCGGAGGACCATTTGAC 452
Qy	181 ATCGTCTCTGAGATGAGATCTGTGAGACGGTGGCTCTGGAAGTCTCAGGCAGAGATAT 240
Db	453 ATCGTCTCTGAGATGAGATCTGTGAGACGGTGGCTCTGGAAGTCTCAGGCAGAGATAT 512
Qy	241 AACATCCCAAGGCTCAAGTGGCCAAAGTCTCCATGAGGTGCTGTATGAGGCGCTGAGC 300
Db	513 AACATCCCAAGGCTCAAGTGGCCAAAGTCTCCATGAGGTGCTGTATGAGGCGCTGAGC 572
Qy	301 AAGGAGAAACAGAGAACTGTTGTTACCTGTGGAAACCTCGAGAGGGGCTTCCCAT - 359
Db	573 A--GAAAAACAAAGAGACTGCTGTTGTTACCTGTGGAAACCTCGAGAGGGGCTTTCCTC 630
Qy	360 CCGGAGAGCCAGACCCAGAGAGGCT-CTTACTCTCTGT-CAGTCCGCTCAAGCGGCC 416
Db	631 CCGGAGAGACCAAGCCAGAGAGGCTCTCAACTCTCTGTGTGTGTCGCTTTCAGCCGCC 689

[illegible]

Db 721 TCCCGAGAGACGACCGAGAGAGCTCTTA 752

RESULT 10
AL844311 597 bp mRNA linear EST 30-JUL-2002
LOCUS AL844311 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AL844311
VERSION AL844311.1 GI:22019093
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 597)
Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and
Sheridan, E.
AUTHORS
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquerry@sanger.ac.uk
Sanger Centre name: scdd10816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool YT_11b v SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.

FEATURES

source

1. 597
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_11b="pool_YT_11b_v_SPD"
/note="Organ: breast; Vector: pZERO-1; Site: 1: Sph1;
Site 2: Sph1; Ductal carcinoma in situ, high-grade, comedo
from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."

BASE COUNT 130 a 186 c 157 g 124 t
ORIGIN

Query Match 40.8%; Score 319.4; DB 9; Length 597;
Best Local Similarity 99.7%; Pred. No. 2.8e-71;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 463 AATGGCTGCTGATATCTACCCGCTCACCCTCCCTCACCCTCCAGGCGCTGCTGAC 522
DB 13 AATGGCTGCTGATATCTACCCGCTCACCCTCCCTCACCCTCCAGGCGCTGCTGAC 72
QY 523 CATTACTGAGCTGGCGATGATCTGCTGCTCACTCAAGAGCCCTGCTCTGAC 582
DB 73 CATTACTGAGCTGGCGATGATCTGCTGCTCACTCAAGAGCCCTGCTCTGAC 122
QY 583 AGGCGTGGCCCTCCCTGCGAAGAGATATACCTCACTGAGCTGAGAGAGACCA 642
DB 133 AGGCGTGGCCCTCCCTGCGAAGAGATATACCTCACTGAGCTGAGAGAGACCA 192
QY 643 CTCAACTGAAAAGCTGAGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAG 702
DB 193 CTCAACTGAAAAGCTGAGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAG 252
QY 703 TCTCTCTCAGTAGAGGCTCCGGAGGCTCCCTCACTTCAATCAAGCTGAATGAGAG 762
DB 253 TCTCTCTCAGTAGAGGCTCCGGAGGCTCCCTCACTTCAATCAAGCTGAATGAGAG 312
QY 763 GCTGCTCTTTGGATGATGCC 783
DB 313 GCTGCTCTTTGGATGATGCC 313

RESULT 11

AL844307 614 bp mRNA linear EST 30-JUL-2002
LOCUS AL844307 pool_AK_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AL844307
VERSION AL844307.1 GI:22019089
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 614)
Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and
Sheridan, E.
AUTHORS
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquerry@sanger.ac.uk
Sanger Centre name: scdd10816.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool AK_11b v SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.

FEATURES

source

1. 614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone_11b="pool_AK_11b_v_SPD"
/note="Organ: breast; Vector: pZERO-1; Site: 1: Sph1;
Site 2: Sph1; Ductal carcinoma in situ, high-grade, comedo
from 41 yo female. Library constructed in the laboratory
of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."

BASE COUNT 134 a 188 c 164 g 128 t
ORIGIN

Query Match 40.2%; Score 314.4; DB 9; Length 614;
Best Local Similarity 99.7%; Pred. No. 5.5e-70;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 468 CTGCTGATACATCTACCGGCTTACCTTCCCTCACTCCAGGCGCTGATGACATTA 527
DB 1 CTGCTGATACATCTACCGGCTTACCTTCCCTCACTCCAGGCGCTGATGACATTA 60
QY 528 CTCTAGCTGGCGATATCTGCTGCTCAAGAGCCCTGCTGAGAGGGC 587
DB 61 CTCTAGCTGGCGATATCTGCTGCTCAAGAGCCCTGCTGAGAGGGC 120
QY 588 TGGCCGCTCCCTGCGAAGAGATATACCTCACTGAGCTGAGAGAGACCACTCA 647
DB 121 TGGCCGCTCCCTGCGAAGAGATATACCTCACTGAGCTGAGAGAGACCACTCA 180
QY 648 CTGAAAAGCTGAGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGCTCT 707
DB 181 CTGAAAAGCTGAGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGCTCT 240
QY 708 TCTCAGTAGAGGCTCCGGAGGCTCCCTCACTTCAATCAAGCTGAATGAGAGCTGT 767
DB 241 TCTCAGTAGAGGCTCCGGAGGCTCCCTCACTTCAATCAAGCTGAATGAGAGCTGT 300
QY 768 CTCTTTGGATGATGCC 783
DB 301 CTCTTTGGATGATGCC 316

RESULT 12

RESULT 15	
BB619854	
LOCUS	627 bp mRNA linear EST 31-AUG-2001
DEFINITION	BB619854 RIKEN full-length enriched, adult male thymus Mus musculus
	CDNA clone 5830437K10.5, mRNA sequence.
ACCESSION	BB619854
VERSION	BB619854.1 GI:15396929
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

